

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт Космических и информационных технологий
институт
Информационные системы
кафедра

УТВЕРЖДАЮ
Заведующий кафедрой ИС
_____ Виденин С.А.
подпись инициалы, фамилия
« ____ » _____ 20 ____ г.

БАКАЛАВРСКАЯ РАБОТА

09.03.02 - Информационные системы и технологии

Программный модуль системы ПО ЭММ для площадного корреляционного
анализа геолого-геофизических данных

Руководитель	_____	доцент, к.т.н.	Н.В. Молокова
	подпись, дата		
Выпускник	_____		Д.А. Белоусова
	подпись, дата		
Консультант	_____	доцент, к.г.-м.н.	С.М.Макеев
	подпись, дата		
Нормоконтролер	_____		Ю.В. Шмагрис
	подпись, дата		

Красноярск 2017

Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт Космических и Информационных Технологий
институт
Информационные Системы
кафедра

УТВЕРЖДАЮ
Зав. кафедрой ИС
_____ С.А.Виденин
подпись инициалы, фамилия
« _____ » _____ 2016 г.

**ЗАДАНИЕ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ
в форме бакалаврской работы**

фамилия, имя, отчество

номер

КОД

наименование

геофизических данных

Утверждена приказом по университету № 2517/с от 01.03.2017г.

«Информационные системы» ИКИТ СФУ

инициалы, фамилия, должность, ученое звание и место работы

исследования земной коры», Бродовой В.В. «Справочник геофизика».

Экспериментальных данных, заключение, список использованных источников.

Office PowerPoint 2010.

Руководитель ВКР

ПОДПИСЬ

Н.В. Молокова

инициалы и фамилия

Задание принял к исполнению

ПОДПИСЬ

Д.А. Белоусова

инициалы и фамилия студента

« » 2016 г.

РЕФЕРАТ

Бакалаврская работа по теме «Программный модуль системы ПО ЭММ для площадного корреляционного анализа геолого-геофизических данных» содержит 50 страниц текстового документа, 22 рисунка, 18 использованных источников.

Объект аудита: комплексирование геофизических методов на этапе предварительного анализа полевых данных.

Цель аудита: разработка программного модуля корреляции свойств независимых друг от друга полей по разведочным данным и выявление участков, пригодных для дальнейшего исследования с целью добычи руды

В результате проведения аудита выявлена корреляция областей сжатия и растяжения слоев геологической среды на исследуемом участке.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 Постановка задачи.....	7
1.1 Связь геофизических признаков и параметров горных пород геологического разреза	9
1.3 Корреляционный анализ данных и расчет выборочных характеристик в комплексировании геофизических методов.....	12
1.4 Обзор подобных решений на современном рынке.....	13
1.5 Выводы по главе 1.....	14
2 Формализация и выбор метода решения	16
2.1 Обоснование математического аппарата.....	17
2.2 Расчет выборочных характеристик	18
2.3 Корреляционный анализ данных.....	19
2.4 Перспективы расширения функционала алгоритма.....	21
2.5 Выводы по главе 2.....	21
3 Разработка программного модуля	22
3.1 Требования.....	22
3.2 Выбор средств	24
3.3 Проектирование	25
3.4 Реализация алгоритма.....	28
3.5 Выбор ограничений на типы входных и выходных данных	34
3.5 Разработка графического интерфейса	37
3.6 Связь с другими модулями в системе	38
3.7 Проверка адекватности результата	38
3.8 Вывод по главе 3	45
4 Анализ экспериментальных данных	46
4.1 Теоретическое обоснование.....	46
4.2 Результаты эксперимента.....	47
ЗАКЛЮЧЕНИЕ	498
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	49

ВВЕДЕНИЕ

Одним из самых востребованных прикладных направлений в современной науке является сложная естественная наука геофизика, которая исследует строение Земли при помощи физических методов. Одним из ее разделов является разведочная геофизика, которая посвящена изучению строения Земли с целью поиска, уточнения залежей полезных ископаемых и выявлению предпосылок для их образования [1]. Разведочная геофизика занимается процессами и объектами, недоступными для прямого исследования, определяя методы, позволяющие, не нарушая структуры среды, исследовать недра Земли и сделать предположения с минимальной степенью ошибки об их содержимом. Этот раздел делится еще на несколько подразделов: сейсморазведка, гравиразведка, магниторазведка и электроразведка.

Исследование геологической среды заключается в исследовании её электромагнитного поля, наблюдая за изменениями которого можно сделать заключение о характере и строении залежей пород. Такие исследования проводятся с помощью специальной аппаратуры, которая на сегодняшний день производится повсеместно – как в России, так и за рубежом. Одним из самых ярких отечественных представителей является аппаратура электромагнитных методов (комплекс ПО ЭММ), начало разработки которой было положено в 60-х годах прошлого века доктором наук Г.Ф. Игнатьевым. Данная аппаратура и ее сопровождение востребованы на рынке для решения геолого-разведочных задач при разведке новых месторождений или строительных работах, так как предлагают возможность исследования земной поверхности без нарушения ее структуры, что значительно снижает стоимость проведения исследовательских работ. На сегодняшний день под руководством Г.Я. Шайдурова, В.И. Иголкина, О.А. Тренина, М.Ф. Хохлова на базе компании «ООО R.D.SCIENCE» в сотрудничестве с СФУ осуществляется поддержка аппаратуры и проведение полевых работ. Сама аппаратура ЭММ изготовлена в ОАО «Алмаззолотоавтоматика» совместно с ООО «Енисейзолотоавтоматика».

Для оптимизации процесса интерпретации данных и ускорении его проведения в разы, а также для обеспечения достойной конкуренции отечественного продукта на рынке на базе кафедры ИС Института Космических и Информационных Технологий СФУ разрабатывается программное обеспечение аппаратного комплекса ЭММ.

Программный комплекс ЭММ состоит из блоков модулей, каждый из которых реализует тот или иной метод исследования земных недр. В данной работе речь пойдет о разработке модуля для площадного корреляционного анализа геолого-геофизических данных.

Наличие автоматизированного программного обеспечения для произведения расчета и обработки геофизических данных, уменьшающего временные и денежные затраты, - одна из очень актуальных проблем современной разведочной геофизики.

В зависимости от рода деятельности и применяемых разведывательных методов (электроразведка, сейсморазведка, химическая разведка и другие), каждая организация определяет собственные узкоспециализированные цели и задачи. Поэтому для анализа полевых данных разведочной геофизики отсутствует единое и легкодоступное для рядовых организациями программное обеспечение. Существующие же на сегодняшний день программные комплексы в основном ориентированы на конкретные предприятия или организации, а так же требуют больших финансовых затрат для их приобретения [2].

Следовательно, на сегодняшний день проблема определения корреляционной связи между параметрами геологической среды геологического разреза действительно актуальна.

Из вышеизложенного следует, объектом исследования является комплексирование геофизических методов, а предметом исследования - прикладные аспекты моделирования корреляции свойств геологической среды.

Целью работы является разработка программного модуля корреляции свойств независимых друг от друга полей по разведочным данным и выявление участков, пригодных для дальнейшего исследования с целью добычи руды [3].

Для достижения цели поставлены следующие задачи:

1. анализ предметной области, в рамках которого изучаются закономерности соответствия:

1.1) геофизических признаков и параметров горных пород исследуемого геологического разреза,

1.2) критериев рациональности выбора методов для последующего комплексирования;

2. построение математической модели физического процесса с помощью методов математической статистики:

2.1) расчет выборочных характеристик,

2.2) корреляционный анализ данных;

3. создание программного модуля

3.1) анализ требований и составление технического задания на разработку модуля,

3.2) выбор средств разработки,

3.3) проектирование программного модуля.

В первой главе в рамках постановки задачи будет исследована предметная область и рассмотрены решения, которые предлагает современный рынок. Во второй главе будут сформированы правила для математического аппарата и математическая модель. В третьей главе будут описаны и обоснованы требования к модулю, средства разработки и непосредственно процесс разработки, а также проведена проверка работы модуля на корректность математической модели. В четвертой главе будет описана непосредственно реализация вычислительного эксперимента. В заключении приведены выводы и сформулировано дальнейшее развитие этой темы.

1 Постановка задачи

Геофизика - комплекс наук, исследующих физическими методами строение Земли, ее физические свойства и процессы, происходящие в ее оболочках.

Разведочная (прикладная) геофизика - направление геофизики, основанное на изучении внутреннего строения Земли, в основном для поиска и уточнения строения залежей полезных ископаемых, а также выявления предпосылок для их образования [1].

Геофизические методы исследований - это научно-прикладной раздел геофизики, предназначенный для изучения верхних слоев Земли, поисков и разведки полезных ископаемых, инженерно-геологических, гидрогеологических, мерзлотно-гляциологических и других изысканий и основанный на изучении естественных и искусственных полей Земли [4].

Комплексирование геофизических методов, по определению, - это геологически и экономически обоснованное сочетание геофизических методов поиска и разведки месторождений полезных ископаемых и сопровождающих их проверочных и оценочных геологических и геохимических видов исследований, применяемых одновременно или в определённой последовательности [5].

Геокартирование (геологическая съемка) - одна из геологических дисциплин, изучающая методы составления геологических карт и их практическое применение.

Повышение точности исследования и определения объектов геологической среды является одной из важнейших задач геокартирования, а также поисков и разведки полезных ископаемых. Для изучения недр были разработаны десятки геологических, инженерно-геологических, гидрогеологических и геоэкологических методов, большинство из которых являются специфичными для решения профильных задач. Конечно, такое их количество свидетельствует об отсутствии универсальных методов. На

практике для получения универсальных аналитических заключений применяется комплексирование геофизических методов.

Применение комплексирования геофизических методов обеспечивает эффективное решение поставленной задачи применительно к определённым либо типовым условиям объекта, относящегося к конкретному классу и типу. Комплексирование может быть применено на любом этапе геологоразведочного процесса, в данном случае этим этапом является стадия геокартирования.

Геокартирование представляет собой этап обработки полевых данных с помощью специальных ГИС-пакетов (например, программа Surfer) с целью построения площадной карты исследованной местности, отражающей фактическое распределение аномальных зон показателей приборов. Например, на рисунке 1 представлена карта распределения результатов геомагнитного исследования зон сжатия земной коры.

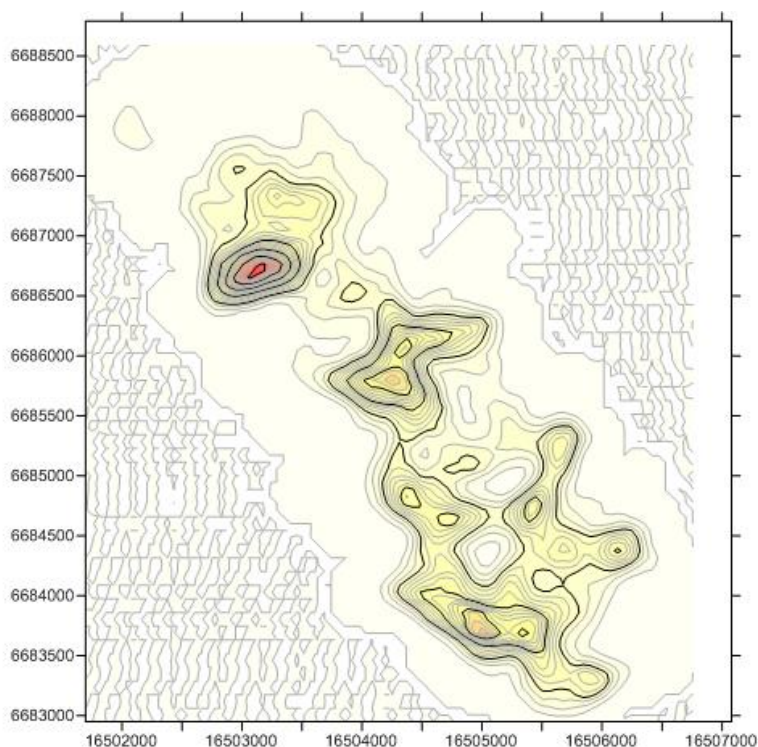


Рисунок 1 - Карта распределения результатов геомагнитного исследования зон сжатия земной коры

1.1 Связь геофизических признаков и параметров горных пород геологического разреза

Для корректного определения связи между исследуемыми параметрами горных пород строится физико-геологическая модель исследуемого объекта и исследуется методический опыт работ при решении аналогичных задач в сходных геологических условиях. Физико-геологическая модель представляет собой обобщённое и формализованное, а также максимально близкое к реальным условиям, представление о геологических и геофизических характеристиках изучаемого объекта и геологической среды. Модели классифицируются по следующим признакам: размер, стадия изучения, геологические и геофизические характеристики и классы рудных объектов (провинции, пояса (области), районы, поля, месторождения и залежи).

Комплексирование применяется для неограниченного количества параметров. Так, комплексирование для двух-трёх и более физических параметров, например, намагниченность и плотность для магнетитовых руд, позволяет снизить влияние неоднозначности решения обратной задачи разведочной геофизики. Также ярким примером может стать анализ электропроводности, поляризуемости, плотности, реже намагниченности, для сульфидных руд.

Сочетание прямых и косвенных геологических признаков (характеристики непосредственно породы и параметры внешней среды), полученные различными методами в одной и той же точке пространства, является одной из главных задач комплексирования. В качестве примера можно привести исследование связи свойств рельефа и электромагнитного поля Земли по разведочным данным. Поскольку не всегда причиной аномальных зон показателей приборов является наличие залежей полезных руд, они также могут возникать из-за влияния рельефа: например, полевые съемки проводятся в гористой местности и холмы или низменности вносят свой вклад в

погрешность полевых измерений, или же в зоне исследований выявлено обилие подземных вод. И для того, чтобы исключить зоны, не имеющие залежи руд, применяются методы корреляционного анализа связи между влиянием рельефа на показания приборов.

1.2 Критерии рациональности выбора методов для последующего комплексирования

Решение тех или иных геофизических задач с помощью комплексирования требуют проведения предварительного обоснования выбранного сочетания методов исследования геологической среды. Обоснование выбора бывает двух видов: геологическое и экономическое.

Геологическое обоснование включает анализ

- характера решаемых задач,
- условий проведения работ,
- характера размещения полезных ископаемых, а также качества и глубины залегания полезных ископаемых.
- требований промышленности к запасам [6].

Экономическое же обоснование подразумевает анализ оптимальности затрат на выполнение геофизических исследований, и их влияние на экономику всего комплекса геологоразведочных работ. В итоге, эффективным считается такое решение задачи, которое удовлетворяет геологическому заключению и решению о целесообразности проведения геологоразведочных работ последующей стадии, локализации перспективной площади, прогнозе размещения и оценке перспектив полезных ископаемых [5].

Рациональным выбором для геофизического комплексирования является такой выбор комплекса методов, который сможет обеспечить однозначное решение поставленной геологической задачи, с минимальной погрешностью в определении местоположения и расшифровке их физических свойств. Самыми популярными методологическими приемами являются:

- переход от мелких масштабов к более крупным, например, осуществляя плавный переход от изучения больших площадей методом попланшетного картирования к разведке перспективных участков,
- переход от сравнительно быстрых (аэрокосмических, морских) к более медленным детальным полевым и подземно-скважинным методам,
- повторение съемок более точной аппаратурой по более густой сети наблюдений,
- переход от интерпретации данных каждого отдельного геофизического метода к комплексной компьютерной обработке всех материалов,
- переход от геологически качественного истолкования материалов к количественному, применяя петрофизические данные.

Исходя из выбора методик, выделяют следующие виды геофизических комплексов:

- типовые комплексы (сочетание избыточного числа методов, на данной стадии исследований обеспечивающее точное решение поставленных задач),
- рациональный комплекс (экономически обоснованный и ограниченный набор типовых комплексов, обеспечивающих надежное решение поставленных задач);
- технологические комплексы (объединение методов по месту и уровню проведения работ): космические, аэрогеофизические, полевые, акваториальные, подземные и скважинные [4].

Проблема разработки теории и методологии комплексирования решается на основе разнообразных информационно-компьютерных технологий, в каждом отдельно взятом частном случае ориентированно на конкретные методы и методики исследования геологической среды.

1.3 Корреляционный анализ данных и расчет выборочных характеристик в комплексировании геофизических методов

Методы математической статистики широко применяются в геофизических исследованиях: для оценки, анализа и прогнозирования результатов геологических работ.

В случае же комплексирования математических методов широко применяются кластерный анализ и корреляций анализ. Первый целесообразно применять, когда необходимо проанализировать результаты более, чем двух методик. Если же сравнение производится по двум показателям (например, связь рельефа и магнитных свойств среды, наложение зон сжатия и расширения земной коры) и имеется два дискретных набора полевых данных, то достаточно провести корреляционный анализ, который покажет силу связи между показателями на тех или иных участках.

В рамках программно-аппаратного комплекса ПО ЭММ комплексирование применяется исключительно на стадии предварительного анализа полевых данных, в частности, как уже указывалось, на этапе картирования. Это позволяет исследователю быстро и надежно определить, какие из участков, на которых были замечены повышенные всплески показателей прибора, являются ложными аномалиями, а какие - истинными. Зачастую, аномалеобразующим фактором может стать не только наличие залежей полезных ископаемых, но и, например, подземная воздушная полость или поземное течение. На практике исследователь дает качественную оценку результатам комплексирования: если корреляция минимальна, а показатели приборов на данном участке скачкообразны, то делается заключение о наличии в данном районе возможных залежей руд и проводятся соответствующие дополнительные исследования, например, геохимические или электрофизические.

Расчет выборочных характеристик применяется для выявления слабых аномалий также при предварительном анализе – сглаживая незначительные

скачки с помощью усреднения показателей приборов относительно географических координат, исследователь получает более полную картину происходящего.

1.4 Обзор подобных решений на современном рынке

Основные конкуренты данной разработки как программного комплекса на отечественном рынке - следующие геофизические комплексы:

- Георадар (этот продукт выпускается многими компаниями, как отечественными, так и зарубежными),
- радар Немфис,
- Новосибирская научная разработка - комплекс ЭМС.

Комплекс ЭММ значительно отличается от этих разработок по следующим показателям:

- Глубина зондирования,
- Устойчивость работы в неблагоприятных условиях (пески, вечная мерзлота, вода, болотистая местность),
- Экономические и рыночные преимущества аппаратной составляющей(цена, наличие ПО интерпретации, стоимость ПО)

Комплекс ЭММ предлагает, в первую очередь, более глубокое зондирование (порядка сотен метров), чем основные конкуренты (до 20м) за счет высокочастотной области методов порядка килогерц, тогда как приборы, предназначенные для нефтяных исследований, работают в области порядка герц, а осуществляющие поверхностные исследования (радары и георадары) в области порядка мегагерц. Данный аппаратный комплекс также дает возможность исследований в самых неблагоприятных условиях (пески, вечная мерзлота, вода и пр.), поскольку не требует, в отличие от основных конкурентов, специальных установок, фиксирующих его на поверхности (исследователю достаточно нести его в руках).

Отсюда следует, что ввиду методических особенностей и соответствующей аппаратной реализации, комплекс ЭММ требует собственного уникального программного обеспечения, в том числе и для предварительного анализа, включающего и комплексирование специфичных методов. Ранее для аппаратуры ЭММ не создавалось специального программного обеспечения, обеспечивающего математическую обработку комплексирования двух методов с помощью корреляционного анализа.

1.5 Выводы по главе 1

В связи с вышеперечисленным можно выделить объект исследования - комплексирование геофизических методов, предмет исследования - прикладные аспекты моделирования корреляции свойств рельефа и электромагнитного поля Земли.

Исходя из предмета исследования, ставится цель данной выпускной квалификационной работы — разработка программного модуля корреляции свойств независимых друг от друга полей по разведочным данным и выявление участков, пригодных для дальнейшего исследования с целью добычи руды.

Чтобы достигнуть поставленной цели, необходимо выполнить ряд задач:

1. анализ предметной области, в рамках которого изучаются закономерности соответствия:

1.1) геофизических признаков и параметров горных пород исследуемого геологического разреза,

1.2) критериев рациональности выбора методов для последующего комплексирования;

2. построение математической модели физического процесса с помощью методов математической статистики:

2.1) расчет выборочных характеристик,

2.2) корреляционный анализ данных;

3. создание программного модуля

- 3.1) анализ требований и составление технического задания на разработку модуля,
- 3.2) выбор средств разработки,
- 3.3) проектирование программного модуля.

2 Формализация и выбор метода решения

Расчет в разработанном программном модуле производится исключительно с помощью формул математической статистики, поскольку связи между литологическим типом руд и их физическими свойствами, а значит и между результатами исследований, проведенных с помощью соответствующих методов, неоднозначны и носят вероятностно-статистический характер.

Разработанный модуль получает в качестве входных данных результат полевых исследований, представляющий собой набор точек X и Y (пример приведен на рисунке 2), каждой из которых соответствует пара показателей магнитных характеристик слоев пород. Этот набор получают путем последовательных замеров с помощью аппаратуры ЭММ на площади с заданным интервалом. Впоследствии результаты замеров располагают на карте с помощью инструментов ГИС-пакета Surfer в виде участков повышенных и пониженных значений электромагнитных показателей.

	A	B	C	D
1	X	Y	FI	Fr
2	16504646	6672170	-0,27469	-0,2209
3	16504686	6672170	-0,27469	-0,2209
4	16504726	6672170	-0,27469	-0,2209
5	16504766	6672170	-0,27469	-0,2209
6	16504806	6672170	-0,27469	-0,2209
7	16504846	6672170	-0,27469	-0,2209
8	16504886	6672170	-0,27469	-0,2209
9	16504926	6672170	-0,27469	-0,2209
10	16504966	6672170	-0,27469	-0,2209
11	16505006	6672170	-0,27469	0,019547
12	16505046	6672170	-0,27469	1,190873
13	16505086	6672170	-0,27469	1,984719
14	16505126	6672170	-0,27469	0,94855
15	16505166	6672170	-0,21572	-0,09861
16	16505206	6672170	-0,12226	-0,2209
17	16505246	6672170	-0,1083	-0,2209
18	16505286	6672170	-0,12853	-0,2209
19	16505326	6672170	-0,15704	-0,2209
20	16505366	6672170	-0,19508	-0,2209

Рисунок 2 – Пример результата полевых исследований двух магнитных характеристик геологической среды

Для проведения корреляционного анализа применяется метод, подобный так называемому «методу скользящего окна»: на заданного размера участках по всей площади карты последовательно с заранее заданным шагом выбираются наборы точек, по которым вычисляется выборочное среднее для географических координат с целью привязать полученный коэффициент корреляции к центру «окна» и коэффициент корреляции для двух наборов результатов полевых исследований.

2.1 Обоснование математического аппарата

С целью формализовать математический аппарат, выбранный для реализации программного модуля, построим математическую модель.

В качестве параметров модели используем следующие:

- (x, y) – пара реальных географических координат точки,
- (z_1, z_2) – пара показателей измерений приборов,
- (\bar{x}, \bar{y}) – пара средних значений по географическим координатам, удовлетворяющих параметрам «окна»,
- a, b – высота и ширина «окна» соответственно (количество дискретных точек),
- r – коэффициент корреляции по показателям приборов, удовлетворяющих параметрам «окна».

Исследование процесса корреляции проведем при следующих условиях:

- Объектом исследования является процесс комплексирования двух параметров геологической среды,
- Показатели этих параметров попарно соответствуют их реальным географическим координатам,
- Наборы показателей параметров геологической среды равны количественно,
- Связь между показателями линейна.

Отсюда, математическая постановка задачи звучит следующим образом:

Считая набор пар $((x, y), (z_1, z_2))$ заданным найти последовательность пар $((\bar{x}, \bar{y}), r)_1, ((\bar{x}, \bar{y}), r)_2, ((\bar{x}, \bar{y}), r)_3 \dots$, удовлетворяющих следующим правилам:

- Наборы для вычисления средних значений и коэффициента корреляции выбираются из упорядоченного набора пар $((x, y), (z_1, z_2))$ пошагово, причем количество элементов равно произведению чисел a и b , количество наборов точек с одинаковым значением координаты y равно длине «окна», а количество наборов точек с одинаковым значением координаты x равно ширине «окна»,
- Для каждого набора рассчитывается выборочное среднее по координате x , выборочное среднее по координате y и коэффициент корреляции между наборами чисел по показателям z_1, z_2 .

Поскольку для любого набора чисел возможно определить его выборочное среднее, и для любых двух наборов, числа из которых сопоставлены друг другу попарно, можно найти коэффициент корреляции, имеем математически корректную задачу.

2.2 Расчет выборочных характеристик

Выборочное наблюдение - это метод наблюдения, при котором обследуется не вся совокупность, а лишь часть ее, отобранная по определенным правилам выборки.

Выборочное наблюдение ставит перед собой задачу - по обследуемой части дать характеристику всей совокупности единиц при условии соблюдения всех правил и принципов проведения статистического наблюдения и научно организованной работы по отбору единиц.

Выборочное (эмпирическое) среднее - это приближение теоретического среднего распределения, основанное на выборке из него [7].

В программе выборочное среднее вычисляется отдельно для показателей географических координат в каждом окне по оси ОХ и оси ОУ по формуле выборочного среднего (формула 2.1).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (2.1)$$

где n – объем выборки, x_i – результат измерения i -ого элемента выборки в окне.

Целью этого процесса является рациональное распределение значения коэффициента корреляции на плоскости: при сглаживании карты по средним значениям, на выходе получается набор точек, фиксирующих центры каждого окна корреляции для сопоставления с коэффициентом корреляции.

2.3 Корреляционный анализ данных

Корреляционный анализ - метод, позволяющий обнаружить зависимость между несколькими случайными величинами.

Корреляция или корреляционная зависимость — статистическая взаимосвязь двух или более случайных величин (либо величин, которые можно с некоторой допустимой степенью точности считать таковыми). При этом изменения значений одной или нескольких из этих величин сопутствуют систематическому изменению значений другой или других величин.

Математической мерой корреляции двух случайных величин служит корреляционное отношение либо коэффициент корреляции. В случае если изменение одной случайной величины не ведёт к закономерному изменению другой случайной величины, но приводит к изменению другой статистической характеристики данной случайной величины, то подобная связь не считается корреляционной, хотя и является статистической [8].

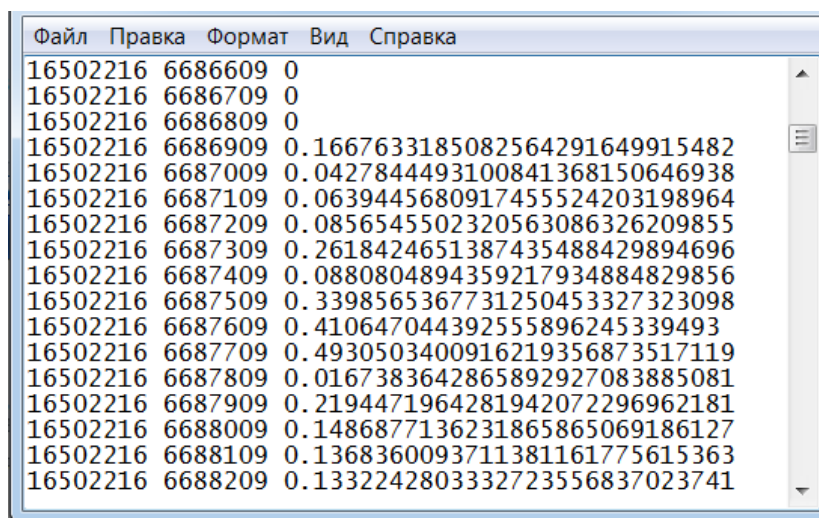
В программном модуле площадного корреляционного анализа для каждого набора данных в каждом окне вычисляется коэффициент корреляции

по показателям, зафиксированным прибором. Известно, что чаще всего корреляционная зависимость между двумя средами носит линейный характер, поэтому расчет производится по классической формуле для двух дискретных наборов чисел (формула 2.2)

$$r(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}}, \quad (2.2)$$

где x_i – элементы из набора показателей, полученных с помощью исследования первым методом, а y_i – элементы из набора показателей, полученных с помощью исследования вторым методом, соответственно.

Результатом работы алгоритма является таблица корреляции, где хранятся как географические координаты, так и «привязанные» к ним значения коэффициента (пример такой таблицы приведен на рисунке 3).



Файл	Правка	Формат	Вид	Справка
16502216	6686609	0		
16502216	6686709	0		
16502216	6686809	0		
16502216	6686909	0.1667633185082564291649915482		
16502216	6687009	0.0427844493100841368150646938		
16502216	6687109	0.0639445680917455524203198964		
16502216	6687209	0.0856545502320563086326209855		
16502216	6687309	0.2618424651387435488429894696		
16502216	6687409	0.0880804894359217934884829856		
16502216	6687509	0.3398565367731250453327323098		
16502216	6687609	0.410647044392555896245339493		
16502216	6687709	0.4930503400916219356873517119		
16502216	6687809	0.0167383642865892927083885081		
16502216	6687909	0.2194471964281942072296962181		
16502216	6688009	0.1486877136231865865069186127		
16502216	6688109	0.1368360093711381161775615363		
16502216	6688209	0.1332242803332723556837023741		

Рисунок 3 – Пример таблицы корреляции

Таблицы корреляции впоследствии используются для генерации карт в программе Surfer, которая наглядно показывает связь между аномалиями, обнаруженными при полевых работах. Получив набор значений коэффициента, программа делает заключение о силе данной связи (чем ближе значение модуля

коэффициента к единице, тем связь сильнее, например, сильной связью считается связь с коэффициентом, значение модуля которого превышает 0,7), отображая это в виде изолиний на карте.

2.4 Перспективы расширения функционала алгоритма

С целью решения большего спектра задач с помощью разработанного модуля планируется расширить функционал в следующих направлениях:

- Создать функцию сопоставления пар по координатам для двух выборок сделанных на одном и том же участке, но в разных точках или с разным шагом,
- Создать функцию автоматического формирования оптимального шага, а также размеров оптимального окна на основании размеров, количества и близости расположения центров наложения сгущения изолиний исследуемых карт,
- Внедрить возможность увеличивать количество входных выборок от трех и более с возможностью комплексирования с помощью анализа средствами нечеткой логики;

2.5 Выводы по главе 2

Итогом этой главы является обоснование выбора математических методов, применяемых для расчетов в разработанном программном модуле и алгоритма анализа входных данных, а также математическое обоснование модели, реализуемой с помощью программных средств.

3 Разработка программного модуля

В данной главе рассматривается непосредственно разработка программного модуля, описываются требования заказчика, приводится обоснование выбора средств и методов разработки, описывается проектирование модуля и его роль в информационной системе, а также подробно разбираются основные этапы разработки.

3.1 Требования

Перед началом разработки для оформления единого видения проекта и его составляющих было проведено интервью с заказчиком и установлены функциональные и нефункциональные требования непосредственно к самой информационной системе, а также к модулю площадного корреляционного анализа.

К функциональным требованиям разрабатываемой информационной системы ПО ЭММ, непосредственно затрагивающим работу модуля площадного корреляционного анализа, относятся следующие требования:

- Каждый модуль системы должен в качестве входных данных принимать файл, содержащий таблицу с подготовленными полевыми данными и их географическими координатами,
- Модуль площадного корреляционного анализа в качестве входных данных получает только один файл, содержащий таблицу из четырех столбцов, первые два из которых – географические координаты, последние два – показатели приборов,
- Каждый модуль системы в качестве выходных данных формировать файл, содержащий таблицу с обработанными данными и их географическими координатами,
- Модуль площадного корреляционного анализа в качестве выходных данных формирует только один файл, содержащий таблицу из трех столбцов,

первые два из которых – географические координаты, последний – значения коэффициента корреляции,

- Каждый модуль в системе должен так или иначе предоставлять пользователю возможность представить полученные данные в графической форме и затем сохранить результат,

- Модуль площадного корреляционного анализа в качестве выходных данных формирует такой файл, который впоследствии можно интерпретировать с помощью ГИС-пакета Surfer,

- Модуль площадного корреляционного анализа должен функционировать внутри семантического блока программ геологического картирования и иметь доступ к файлам других модулей этого блока,

- Доступ пользователя к любому модулю системы не зависит от роли пользователя в системе;

К нефункциональным требованиям разрабатываемой информационной системы ПО ЭММ, непосредственно затрагивающим работу модуля площадного корреляционного анализа, относятся следующие требования:

1. Требования производительности:

- Модули системы ПО ЭММ должны достаточно быстро обрабатывать информацию, так как объем обрабатываемых данных огромен — сотни тысяч строк измерений во входящих файлах,

- Модули системы ПО ЭММ должны позволять загружать и сохранять файлы в наиболее популярных форматах - .csv, .dat, .xls, .xlsx,

- Модуль площадного корреляционного анализа должны позволять загружать файлы в следующих форматах - .csv, .dat,

- Модуль площадного корреляционного анализа должны позволять сохранять файлы в следующих форматах - .dat;

2. Эксплуатационные требования

- Система ПО ЭММ должна работать операционных системах семейства Windows;

3. Требования к пользовательскому интерфейсу:
 - Интерфейс всех модулей системы ПО ЭММ должен быть дружелюбным и интуитивно понятным,
 - Интерфейс всех модулей системы ПО ЭММ должен соответствовать функциональным требованиям к системе;
4. Требования безопасности
 - Нельзя размещать программный код системы на публичных веб-ресурсах

3.2 Выбор средств

В качестве языка разработки был выбран объектно-ориентированный язык C#. Выбор в его пользу был сделан для того, чтобы достичь следующих целей:

- Инкапсуляция функционала модулей внутри блоков и реализаций геофизических методов для исключения вмешательств извне и неправильного использования данных
- Построение семантических объектов в геофизических исследованиях с помощью классов
- Разделение классов, хранящих основной функционал, и классов, хранящих объекты и действия с ними
- Наследование функционала базовых классов для семантических блоков программы в классах внутренних модулей
- Реализация взаимодействия между модулями с помощью паттерна «абстрактная фабрика»

В качестве среды разработки используется интегрированная среда разработки программного обеспечения Microsoft Visual Studio 2015 Community Edition. Данный продукт позволяет разрабатывать как консольные приложения, так и приложения с графическим интерфейсом, в том числе с

поддержкой технологии Windows Forms и WPF, а также веб-сайты, веб-приложения, веб-службы, мобильные приложения [9].

Для оптимизации управления и разработки была применена распределённая система управления версиями Git. В качестве сервиса для хостинга проекта и совместной разработки был выбран веб-сервис Bitbucket. Главным его достоинством является возможность создания закрытого хранилища данных, к которому нет доступа ни у кого, кроме его владельцев, что отвечает требованиям заказчика о сохранности технологий.

3.3 Проектирование

Модуль для расчета корреляции данных создан на основе уже разработанной ранее архитектуры. Данная архитектура была специально разработана для реализации программного комплекса ПО ЭММ, моделирующего процесс расчета и обработки полевых данных разведочной геофизики. Для оптимальной реализации требований заказчика и с целью создать условия для дальнейшего гибкого расширения системы был выбран архитектурный шаблон разработки MVVM.

MVVM (Model-View-ViewModel) - это шаблон, который появился с целью исключения ограничений шаблонов MVC и MVP, и объединяющий их основные сильные стороны. Эта модель впервые появилась в составе фреймворка Small Talk в 1980-х, и была позднее улучшена с учетом обновленной модели презентаций (MVP).

Шаблон MVVM имеет три основных компонента: модель, которая представляет бизнес-логику приложения, представление пользовательского интерфейса XAML, и представление-модель, в котором содержится вся логика построения графического интерфейса и ссылка на модель, поэтому он выступает в качестве модели для представления.

На рисунке 4 представлена диаграмма, которая демонстрирует структуру шаблона MVVM [10].

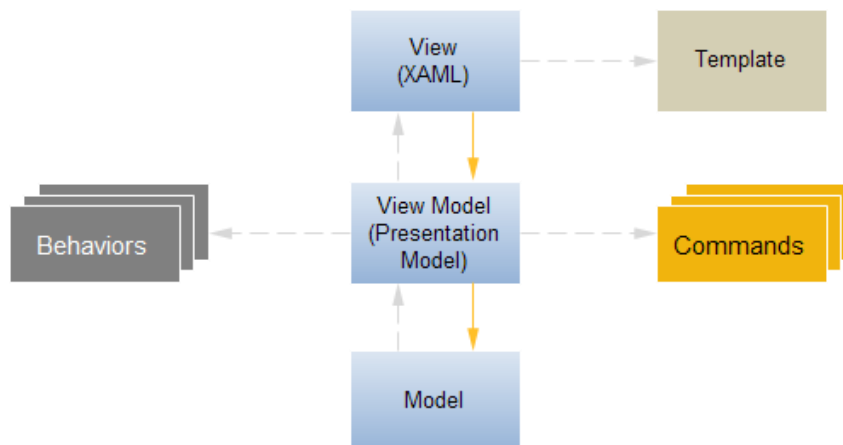


Рисунок 4 - Общая схема шаблона MVVM

На основании требований заказчика к системе разработчиками было принято решение применить сервисно-ориентированный подход, предполагающий наличие в структуре программы независимых сервисов, каждый из которых реализует смысловую функцию обозначенного блока программ.

Смысловая нагрузка каждого из блоков определена следующим образом:

1. Блок базовых программ – хранит все возможные формулы, которые могут быть задействованы для вычислений
2. Блок программ геологического картирования – обработка полевых данных, имеющих жесткую географическую привязку
3. Блок структурно-геометрических программ – обработка геометрических вычислений, а так же построение кривых и графиков на основе полевых данных
4. Блок электрофизических программ - функциональная реализация методов разведочной геофизики
5. Блок программ геологического моделирования – моделирование разрезов на основе полевых данных [11].

Модуль площадного корреляционного анализа относится ко второму семантическому блоку.

В системе ПО ЭММ с целью унификации всех модулей и безопасности их, был применен принцип инверсии управления. Это упрощает расширение возможностей системы, при этом оставляя контроль над потоком управления программы за каркасом. Применяя порождающий шаблон проектирования «абстрактная фабрика», система создает экземпляры модулей на основе Dependency injection фреймворка «Ninject». Это позволяет реализовать независимость абстракции от деталей, так как по определению фабрика предоставляет интерфейс для создания семейств взаимосвязанных или не взаимосвязанных объектов, не специфицируя их конкретных классов.

Для манипулирования шаблоном «инверсия управления» и контролем над объектами в системе используется так называемый паттерн менеджер, необходимый для грамотного управления объектами любого типа (в данной системе такими объектами являются модули блоков). Он обеспечивает своевременное и безопасное получение, удаление или замену модулей по мере необходимости.

В каждый модуль в системе будут входить следующие объекты:

1. Класс или папка `ModuleDataModel`, где хранятся все используемые модулем доменные объекты;
2. Класс `ModuleDataWorker`, где хранятся реализации всех геофизических функций, необходимых для корректной работы модуля;
3. Класс `ModuleViewModel`, реализующий функции взаимодействия пользовательского интерфейса и модели данных и хранящий экземпляры `ModuleDataModel` и `ModuleDataWorker`;
4. Класс `ModuleView`, реализующий пользовательский интерфейс [11].

В работе модуля площадного корреляционного анализа входные данные поступают в `ModuleDataModel`, где происходит преобразование входных данных в площадную матрицу. Полученная матрица потом используется для произведения анализа по формулам, которые прописаны в `ModuleDataWorker` (класс, хранящий реализации геофизических функций). И затем в классе `ModuleViewModel` происходит реализация функций, управляемых

пользователем с помощью элементов графического интерфейса (загрузка и сохранение файлов, пересчет данных). Здесь же предусматривается защита от некорректного ввода данных пользователем: запрещен некорректный ввод размера «окон», в том числе больший размера матрицы, а также символьных строк вместо чисел.

3.4 Реализация алгоритма

В моей работе для каждого функционального элемента в модуле были созданы модель-представление с использованием языка разметки XAML и модель представленная классом. Были разработаны:

- одна модель-представление,
- 3 класса-модели: `DataPoint`, представляющий входные данные и содержащий четыре поля типа `decimal` для географических координат и показателей приборов (листинг этого класса представлен на рисунке 5), `Point3D`, представляющий выходные данные и содержащий три поля типа `decimal` для географических координат и показателя коэффициента корреляции (листинг этого класса представлен на рисунке 5), служебный класс `DataStorage`, соединяющий набор точек в матрицу, предварительно оптимизировав их положение в матрице относительно их реальных географических координат, выходные данные этого класса являются входными данными для класса, хранящего реализации геофизических функций,
- Служебный класс `DataWorker`, предназначенный для обработки матрицы точек и пошаговой выборки из нее наборов данных для вычисления коэффициента корреляции, выходные данные этого класса сохраняются средствами класса `ViewModel` в файл формата `.dat` для дальнейшей обработки средствами ГИС-пакета

В каждом классе модели данных сделано несколько конструкторов для возможности объявления объектов этого класса с разными параметрами – программа предусматривает создание класса с пустыми полями, класса с

полями типа `double` и типа `decimal`, причем при задании полей типа `double`, программа автоматически преобразует их в поля типа `decimal`. Создание конструкторов с разными параметрами отвечает парадигме объектно-ориентированного программирования и реализует такие принципы как полиморфизм и инкапсуляция.

Программа считывает данные из файла по команде, выполнение которой пользователь инициирует с помощью графического интерфейса, в виде неотсортированного списка значений, затем, в классе `DataStorage` проводится проверка на упорядоченность этих данных и относительно поля `Y` формируется массив массивов объектов (так называемая неравномерная матрица).

```
5 public class DataPoint
6 {
7     public DataPoint()
8     {
9     }
10
11     public DataPoint(decimal x, decimal y, decimal z1, decimal z2)
12     {
13         X = x;
14         Y = y;
15         Z1 = z1;
16         Z2 = z2;
17     }
18     public DataPoint(double x, double y, double z1, double z2)
19     {
20         X = (decimal)x;
21         Y = (decimal)y;
22         Z1 = (decimal)z1;
23         Z2 = (decimal)z2;
24     }
25     public decimal X { get; set; }
26     public decimal Y { get; set; }
27     public decimal Z1 { get; set; }
28     public decimal Z2 { get; set; }
29
30 }
31
32 }
```

Рисунок 5 – Листинг класса `DataPoint`

```

3 public class Point3D
4 {
5     public Point3D()
6     {
7
8     }
9
10    public Point3D(decimal x, decimal y, decimal z)
11    {
12        X = x;
13        Y = y;
14        Z = z;
15    }
16
17    public Point3D(double d, double d1, double d2)
18    {
19        X = (decimal) d;
20        Y = (decimal) d1;
21        Z = (decimal) d2;
22    }
23
24    public decimal X { get; set; }
25    public decimal Y { get; set; }
26    public decimal Z { get; set; }
27
28    }
29 }

```

Рисунок 6 - Листинг класса Point3D

Матрица сделана с жесткой привязкой значений для того, чтобы максимально точно отобразить реальное расположение точек на поверхности: необходимо учесть, что в реальных полевых условиях чаще всего нельзя измерить идеальную поверхность (например, исследуется берег реки), и из-за этого массивы в строках смещены относительно друг друга и не равны по длине. Для рационального и эффективного расположения объектов матрицы в процессе формирования делается построчная оптимизация. Функция, формирующая каждую следующую строку, проверяет насколько первый элемент строки близок к элементам предыдущей строки по координате X. Найдя самый близкий элемент, она смещает текущую строку относительно его номера. Все пустым ячейкам присваивается значение null. Ключевое слово null является литералом, представляющим пустую ссылку, которая не ссылается ни на один объект. Для того, чтобы избежать последующих конфликтов при

математических операциях в функциях класса DataWorker предусмотрены проверки на нулевое значение аргумента.

Для оптимизации работы пользователя с программой бы также написан калькулятор значений, позволяющий на основании массива входных данных и расстояния между точками замеров в метрах оценить размер матрицы, которая будет обрабатываться, как в точках, так и в метрах, а также задать размеры окна, соответственно, как в точках, так и в метрах. Это позволит пользователю выбрать оптимальный размер окна и шаг между окнами. Интерфейс калькулятора матрицы представлен на рисунке 7, интерфейс калькулятора окна – на рисунке 8.

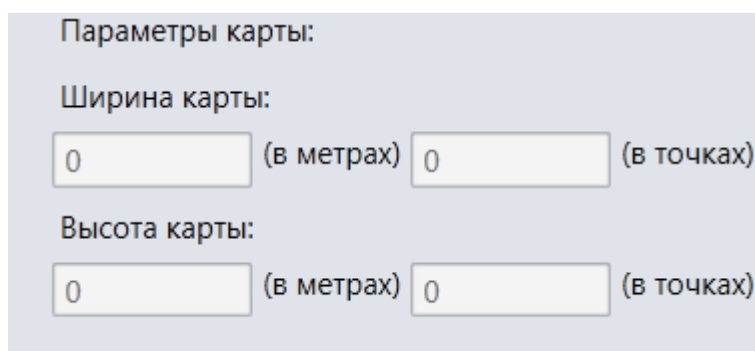


Рисунок 7 - Интерфейс калькулятора матрицы

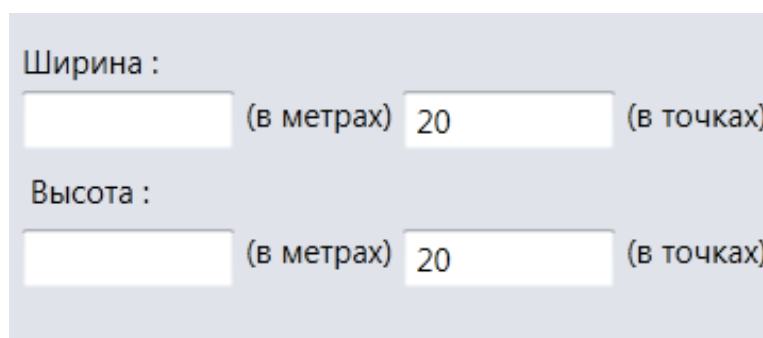


Рисунок 8 - Интерфейс калькулятора окна

После формирования матрицы и передачи её для обработки в класс DataWorker, совершается циклический обход матрицы, на каждом шаге которого выбираются данные, попадающие в окно, заданное с помощью калькулятора окон и для них производятся вычисления выборочного среднего и

коэффициента корреляции. Эти значения записываются в поля нового объекта класса Point3D, который помещается в список значений-результатов. Этот список заполняется до тех пор, пока в матрице есть значения, которые попадают под ограничения окна. После завершения цикла класс DataWorker возвращает классу ViewModel список результатов, который и является таблицей корреляции, где хранятся как географические координаты, так и «привязанные» к ним значения коэффициента. Последующее наложение полученной карты на уже имеющиеся наглядно демонстрирует исследователю-геологу наличие связи между аномалией, обнаруженной прибором, со свойствами рельефа.

После чего таблица записывается в файл, который впоследствии обрабатывается с помощью средств программы Surfer.

На рисунке 9 представлено пересечение карт двух электромагнитных показателей одной и той же области, а на рисунке 10 представлена карта, полученная с помощью модуля корреляционного анализа, выявляющая области пересечения и силу связи между слоями. Сила связи на карте показана с помощью цветовой шкалы - чем больше глубина цвета, тем сильнее связь, в данном случае чем ближе цвет к темно-синему, тем ближе коэффициент корреляции к единице (информация о соответствии цвета на карте показателю коэффициента приведена на цветовой шкале справа).

Как можно заметить, карта, полученная после автоматизации с помощью обработки средствами информационно-компьютерных технологий, визуально воспринимается и читается лучше, что, несомненно, сказывается на качестве анализа, а также его скорости и точности.

Основной задачей программного пакета «Surfer» является расчет значений параметра в узлах регулярной сетки (Grid) по исходным данным в произвольных точках области. Наиболее часто геологи сталкиваются с этой задачей при построении карты поверхности пластов по данным бурения скважин. Интерполяция такого рода - это главная, но далеко не единственная задача, которую можно решить с использованием Surfer.

Существует большое количество программ, решающих подобные задачи. Главным достоинством пакета Surfer является его распространенность, дешевизна, простота в освоении. Исторически сложилось так, что эта программа получила широкое распространение и, по сути, стала стандартом программ подобного типа, что подтверждается включением формата файлов Surfer в программы других производителей.

Цифровая модель поверхности традиционно представляется в виде значений в узлах прямоугольной регулярной сетки, дискретность которой определяется в зависимости от конкретной решаемой задачи. Для хранения таких значений Surfer использует собственные файлы типа GRD (двоичного или текстового формата), которые уже давно стали своеобразным стандартом для пакетов математического моделирования [12].

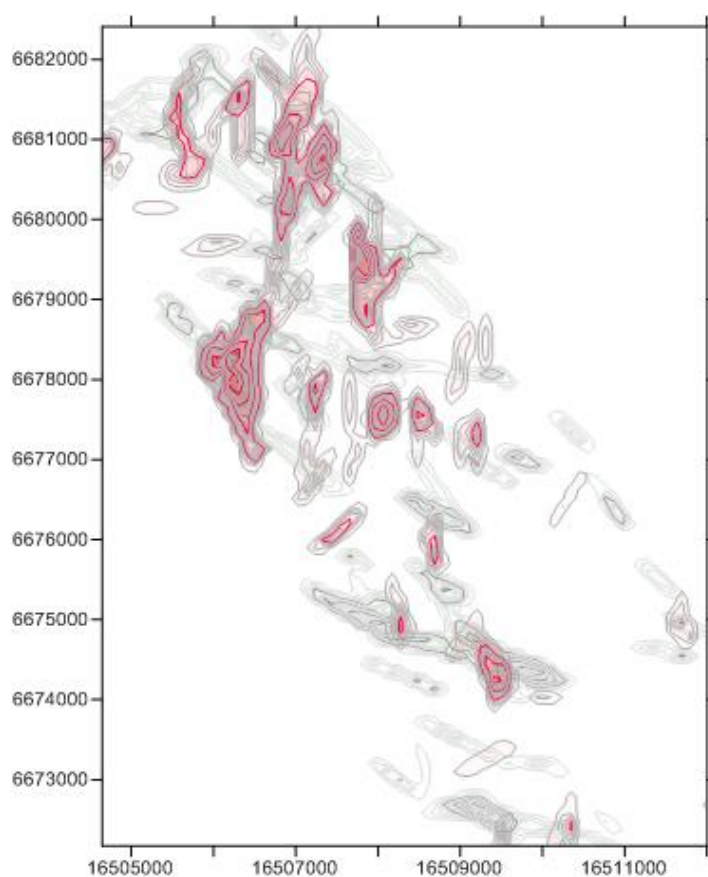


Рисунок 9 - Пересечение карт двух электромагнитных показателей

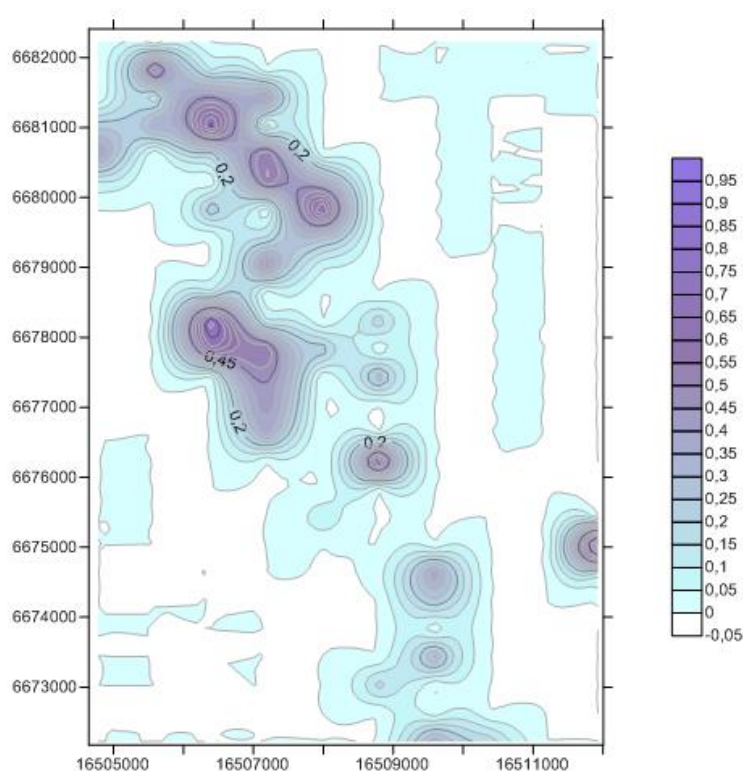


Рисунок 10 – Карта корреляции

3.5 Выбор ограничений на типы входных и выходных данных

Входными параметрами для модуля площадного корреляционного анализа являются файлы следующих форматов:

- .scv,
- .dat;

CSV (от англ. Comma-Separated Values — значения, разделённые запятыми) — текстовый формат, предназначенный для представления табличных данных.

Для данного формата характерны следующие особенности:

- Каждая строка файла — это одна строка таблицы,
- Разделителем (англ. delimiter) значений колонок является символ запятой (,),
- Значения, содержащие зарезервированные символы (двойная кавычка, запятая, точка с запятой, новая строка) обрамляются двойными

кавычками ("). Если значение содержит кавычки — они представляются в файле в виде двух кавычек подряд [13].

Расширение DAT – это универсальный файл данных, который может быть создан и использован множеством приложений. Обычно доступ может предоставляться только приложению, создавшему данный DAT файл. Файл .DAT может содержать данные как в текстовом, так и в двоичном формате. Текстовый DAT файл может быть открыт с помощью текстового редактора, например, Блокнот или Notepad++. Расширение файла DAT также широко используется для некоторых системных файлов в MS-DOS в операционной системе семейства Windows. Очень часто .DAT файлы используют для настроек такие программы как, например, Yahoo! Instant Messenger и Miranda IM [14].

При выборе форматов необходимо было предусмотреть возможные форматы, которые будут поступать в результате дополнительных исследований методами, доступными для комплексирования с методами, применяемыми аппаратурой комплекса ЭММ.

Наиболее популярным форматом является формат .xls, в котором могут поступать данные от исследований от партнерских лабораторий, но на практике было определено, что его обработка занимает очень большое количество времени. При преобразовании данного формата в .csv с разделителем, время, затрачиваемое на обработку данных, сокращается в разы, поэтому было решено отказаться от него и заменить его на .csv.

В пользу выбора формата .csv говорит также тот факт, что выходными данными аппаратной части комплекса ЭММ являются исключительно файлы формата .csv.

Файлы формата .dat генерируются ГИС-пакетом Surfer и их обработка является частью полики связи между модулями, а также оптимизирует процесс обработки данных, полученных не с помощью аппаратуры комплекса ЭММ.

В качестве выходных данных модуль генерирует только файлы формата .dat для оптимальной обработки их в программе Surfer.

На рисунке 11 представлен скриншот файла формата .csv, а на рисунке 12 – формата .dat.

	A	B	C	D	E
1	16504646	6672170	-0,27469	-0,2209	
2	16504686	6672170	-0,27469	-0,2209	
3	16504726	6672170	-0,27469	-0,2209	
4	16504766	6672170	-0,27469	-0,2209	
5	16504806	6672170	-0,27469	-0,2209	
6	16504846	6672170	-0,27469	-0,2209	
7	16504886	6672170	-0,27469	-0,2209	
8	16504926	6672170	-0,27469	-0,2209	
9	16504966	6672170	-0,27469	-0,2209	
10	16505006	6672170	-0,27469	-0,2209	
11	16505046	6672170	-0,27469	-0,2209	
12	16505086	6672170	-0,27469	-0,2209	
13	16505126	6672170	-0,27469	-0,2209	

Рисунок 11 – Входной файл модуля в формате .csv

ID	Column 1	Column 2	Column 3	Column 4
16504646	6672293	-0.27469281091203	-0.22089683999908	
16504666	6672293	-0.27469281045726	-0.22089684142281	
16504686	6672293	-0.27469281086914	-0.22089684175701	
16504706	6672293	-0.27469281058699	-0.22089684150381	
16504726	6672293	-0.27469281029173	-0.22089684126608	
16504746	6672293	-0.27469280999432	-0.22089684104984	
16504766	6672293	-0.27469280985205	-0.2208968408926	
16504786	6672293	-0.27469281231037	-0.22089684292985	
16504806	6672293	-0.27469281148183	-0.22089684214754	
16504826	6672293	-0.27469281058132	-0.22089684132881	
16504846	6672293	-0.27469280972468	-0.22089684056408	
16504866	6672293	-0.2746928102788	-0.22089683956585	
16504886	6672293	-0.27469280864507	-0.2208968400711	
16504906	6672293	-0.27469280869642	-0.22089684023189	
16504926	6672293	-0.27469280876874	-0.22089684040491	
16504946	6672293	-0.27469280882699	-0.22089684056909	
16504966	6672293	-0.27469281279483	-0.22089684266098	
16504986	6672293	-0.27469281257278	-0.22089684352511	
16505006	6672293	-0.2746928121465	-0.22089684312688	
16505026	6672293	-0.2746928116475	-0.22089684266893	
16505046	6672293	-0.27469281108825	-0.22089684215061	
16505066	6672293	-0.27469280948023	-0.22089684100305	
16505086	6672293	-0.27469280727239	-0.22089683915208	
16505106	6672293	-0.27469280763696	-0.22089683942191	
16505126	6672293	-0.27469280806	-0.22089683974206	
16505146	6672293	-0.27469280906054	-0.22089684000572	

Рисунок 12 – Входной файл модуля в формате .dat

3.5 Разработка графического интерфейса

Для разработки графического интерфейса была выбрана технология WPF (Windows Presentation Foundation), являющаяся частью экосистемы платформы .NET.

За счет того, что приложения WPF основаны на DirectX, в отличие от приложений предыдущего поколения (на основе технологии WinForms), основанных на User32 и GDI+, основная часть работы по моделированию графики, независимо от её сложности, ложится на графический процессор на видеокарте, что положительно сказывается на аппаратном ускорении графики.

Одной из важных особенностей WPF является использование языка декларативной разметки интерфейса XAML, основанного на XML. И отсекающему необходимость программного объявления графики и элементов управления [15].

Графический интерфейс главного окна разработанного модуля представлен на рисунке 13.

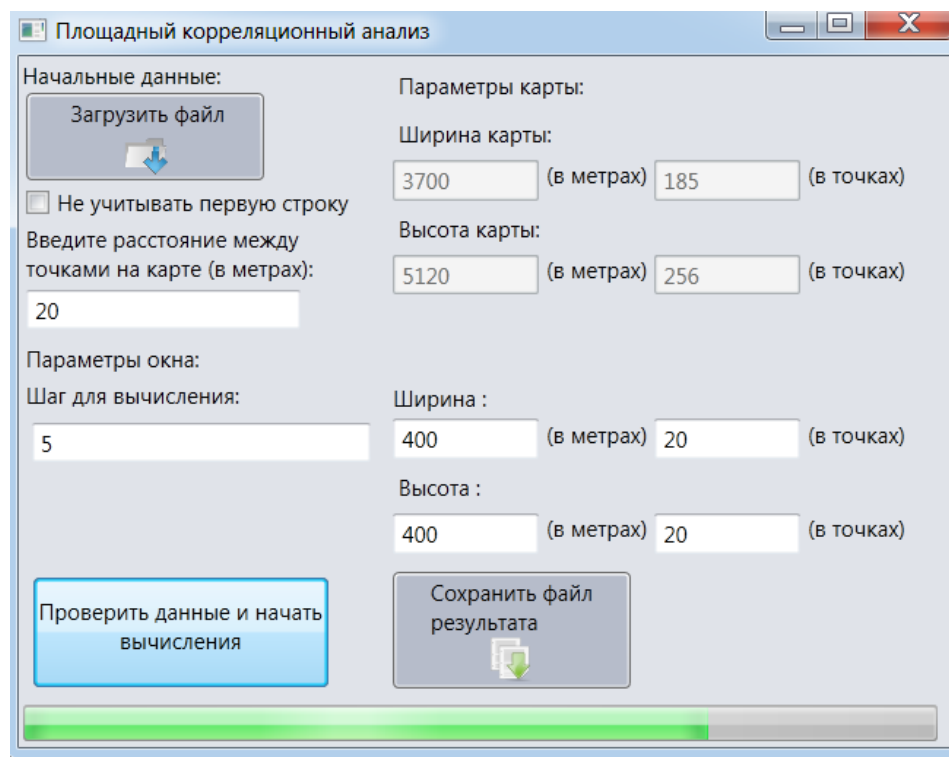


Рисунок 13 – Главное окно приложения

Поскольку приложение рассчитано на использование на различных устройствах разных поколений, преимуществом WPF становится то, что все элементы измеряются в независимых от устройства единицах, и приложения легко масштабируются под разные экраны с разным разрешением.

Приложения, написанные на WPF легко интегрируются с WinForms, благодаря чему, облегчается интеграция новых модулей системы с уже написанными на устаревших технологиях.

3.6 Связь с другими модулями в системе

Связь с другими модулями в рамках программного комплекса ПО ЭММ реализована двумя путями:

- Использование все модулями системы файлов единого формата
- Дизайнерское решение на уровне блока программ геологического картирования, позволяющее задать для одного файла поочередную обработку его разными методами без замены или перезагрузки этого файла

3.7 Проверка адекватности результата

Под адекватностью понимается степень соответствия результатов, полученных по разработанной модели, данным эксперимента или тестовой задачи. Непосредственно перед проверкой на адекватность, проводится предварительная проверка работы выбранных методов решения, комплексная отладка и устраняются все ошибки и конфликты, связанные с программным обеспечением.

Проверка адекватности модели преследует следующие цели:

- убедиться в справедливости совокупности гипотез, сформулированных на этапах концептуальной и математической постановок,
- установить, что точность полученных результатов соответствует точности, оговоренной в техническом задании.

Проверка разработанной математической модели выполняется двумя способами:

- путем сравнения с имеющимися экспериментальными данными о реальном объекте
- путем сравнения с результатами других, созданных ранее и хорошо себя зарекомендовавших моделей.

В первом случае говорят о проверке путем сравнения с экспериментом, во втором — о сравнении с результатами решения тестовой задачи.

Различают качественное и количественное совпадение результатов сравнения. В первом случае требуется наличие совпадений только характерных особенностей в распределении исследуемых параметров: совпадение экстремальных точек, положительное или отрицательное значение параметра в них, возрастание или убывание функции на совпадающих отрезках и т.д. Фактически, при качественном сравнении оценивается совпадение лишь вида функции распределения параметров.

Количественное сравнение допускается только в случае успешного прохождения тестов при качественном сравнении и включает в себя проверку точности дискретных значений функции. При количественном сравнении очень важна проверка точности исходных данных для моделирования и соответствующих им значений сравниваемых параметров [16].

В данной работе качественная проверка результата была выполнена путем сравнения с экспериментом, а количественная — путем сравнения с результатами тестовой задачи.

Для качественной проверки были взяты данные эксперимента, в рамках которого проводилось сравнение данных сдвиговых сжатий и растяжений (левые и правые сдвиги) геологической среды. Проверка проводилась средствами программы Surfer — очевидно, что значение коэффициента корреляции на пересечении изолиний карт должно превышать 0 и, в зависимости от густоты расположения изолиний, приближаться к единице.

На рисунке 14 представлено наложение двух карт, которые исследовались в рамках эксперимента.

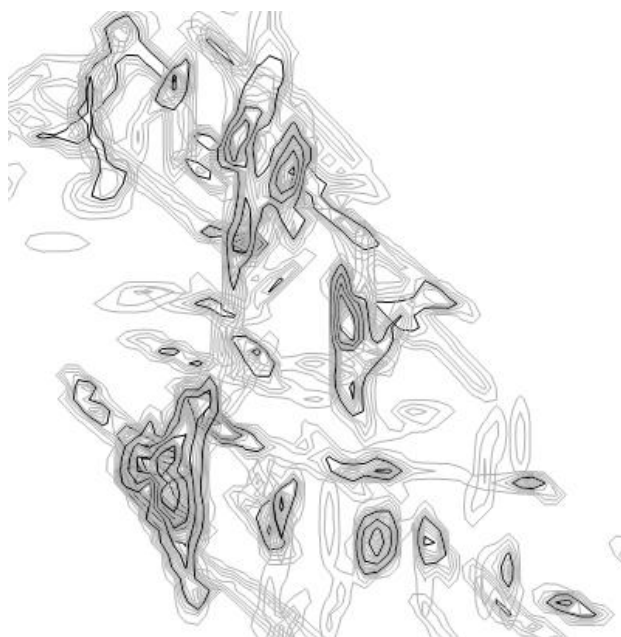


Рисунок 14 – Наложение двух карт

На рисунках 15 и 16 представлены карты проекций по отдельности.

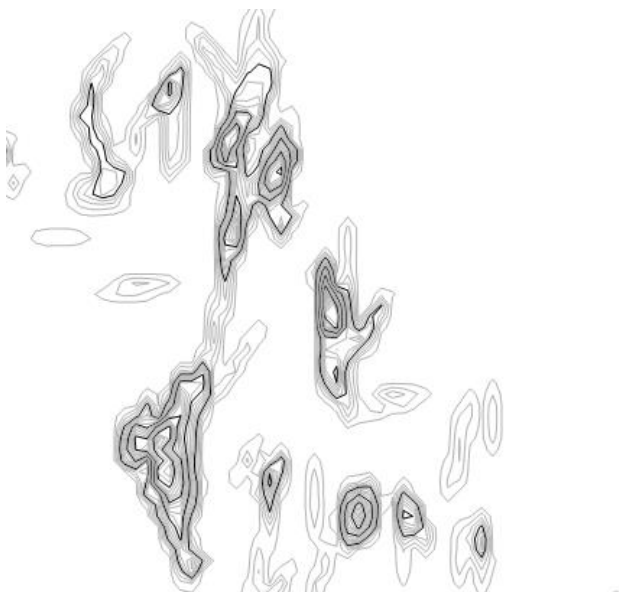


Рисунок 15 – Карта левых сдвигов



Рисунок 16 – Карта правых сдвигов

В рамках эксперимента был проведен корреляционный анализ для окон размером 20, 30 и 35 точек (рисунки 19 – 21 соответственно), причем расстояние между каждой точкой составляет 20 метров. Показателем адекватности модели служит совпадение центров изолиний и самих проекций, и центров корреляции независимо от размера окна, чем шире окно – тем явственнее должны совпасть центры. Решить эту задачу позволяет инструмент пакета Surfer «Point», позволяющий отметить точку на одной карте и, разместив её в отдельном слое, проследить за положением точки относительно изолиний других карт на этой же местности.

Зависимость между аномалиями входных данных представлена на рисунке 17 в виде графика.

На рисунке 18 представлено наложение карт с фиксацией центров сгущения изолиний с помощью инструмента «Point» (крестообразные отметки на карте).

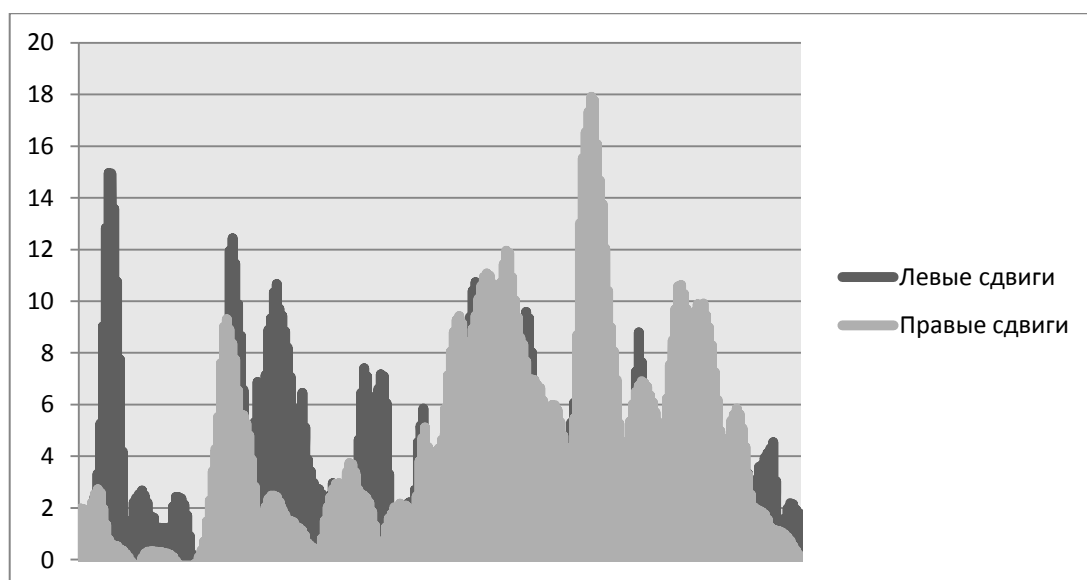


Рисунок 17 – График зависимости входных параметров

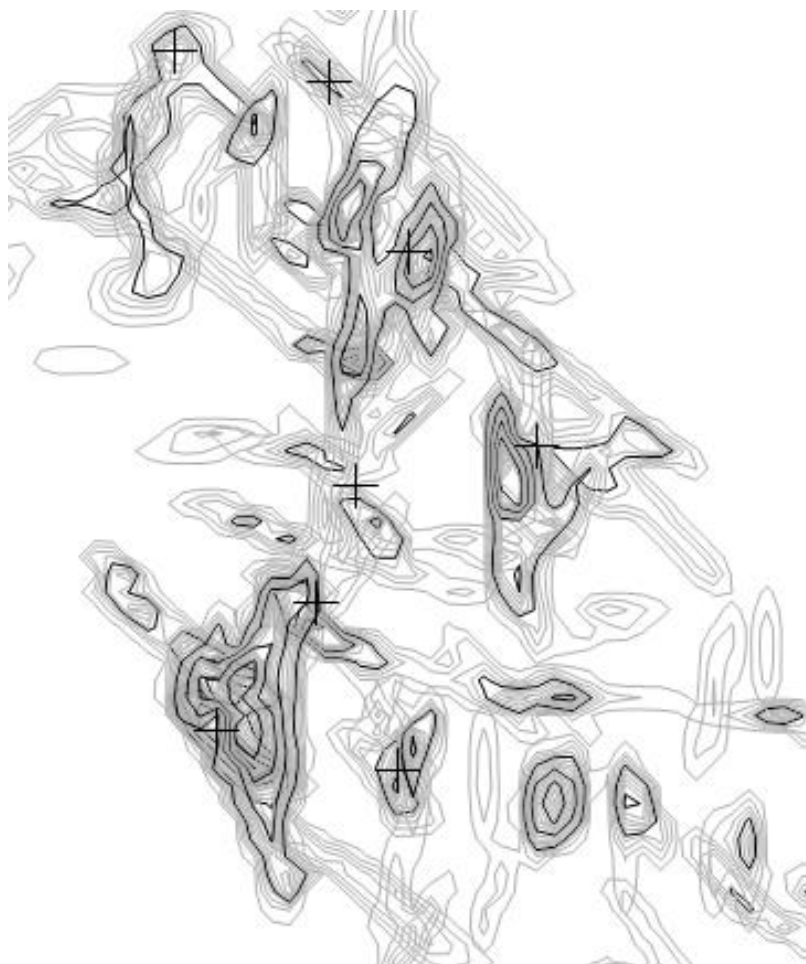


Рисунок 18 – Центры сгущения изолиний

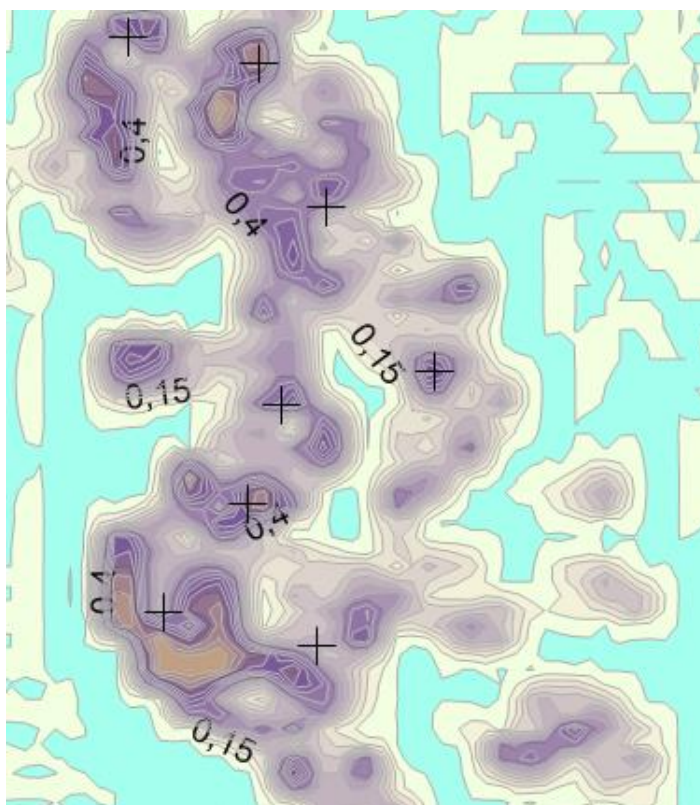


Рисунок 19 – Центры сгущения изолиний, спроецированные на карту корреляции (ширина окна 20)

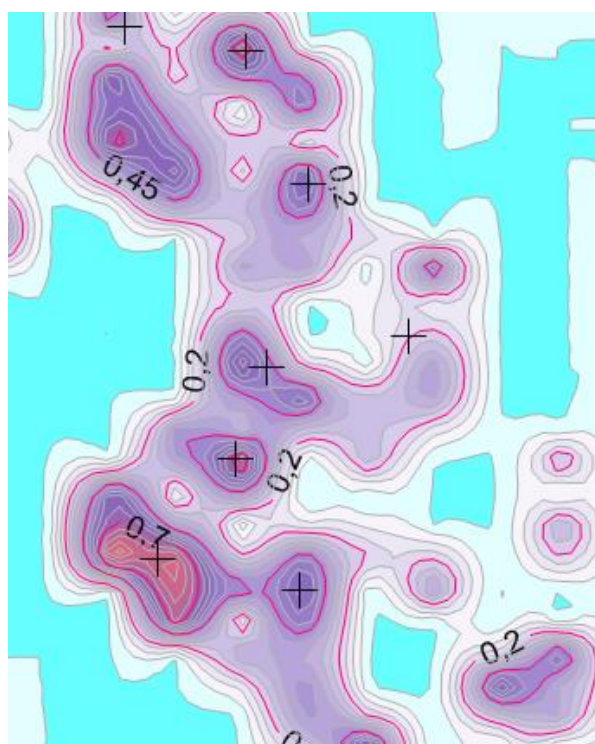


Рисунок 20 - Центры сгущения изолиний, спроецированные на карту корреляции (ширина окна 30)

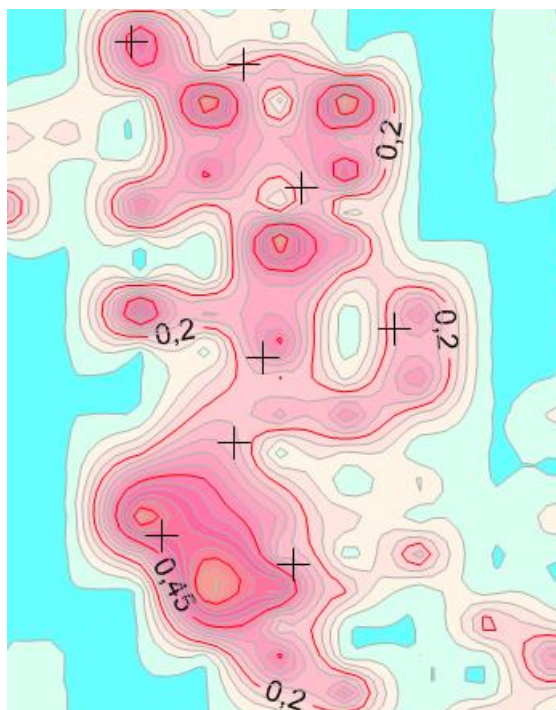


Рисунок 21 - Центры сгущения изолиний, спроецированные на карту корреляции (ширина окна 35)

По результатам количественной оценки расчета коэффициента корреляции был получен коэффициент детерминации равный 0.9862 (98.62%), т.е. точность модели высокая. Остальные 1.38% изменений объясняются факторами, не учтенными в модели (а также ошибками спецификации).

На рисунке 22 представлен график разностей между показателями коэффициента корреляции, полученными с помощью работы модуля, и показателями, рассчитанными вручную в программе Excel с применением функции «КОРРЕЛ» для квадратного окна со стороной 20 точек (расстояние между точками – 20 метров). Как видно из графика, максимальное отклонение не превышает 0,0018, что можно назвать результатом с высокой точностью.

Проверка вычисления модулем выборочного среднего проводилась путем сравнения с показателями, рассчитанными вручную в программе Excel с применением функции «СРЗНАЧ» для квадратного окна со стороной 20 точек (расстояние между точками – 20 метров). В результате количественной

проверки было выявлено полное совпадение значений и получен коэффициент детерминации равный 1 (100%), т.е. точность модели высокая.

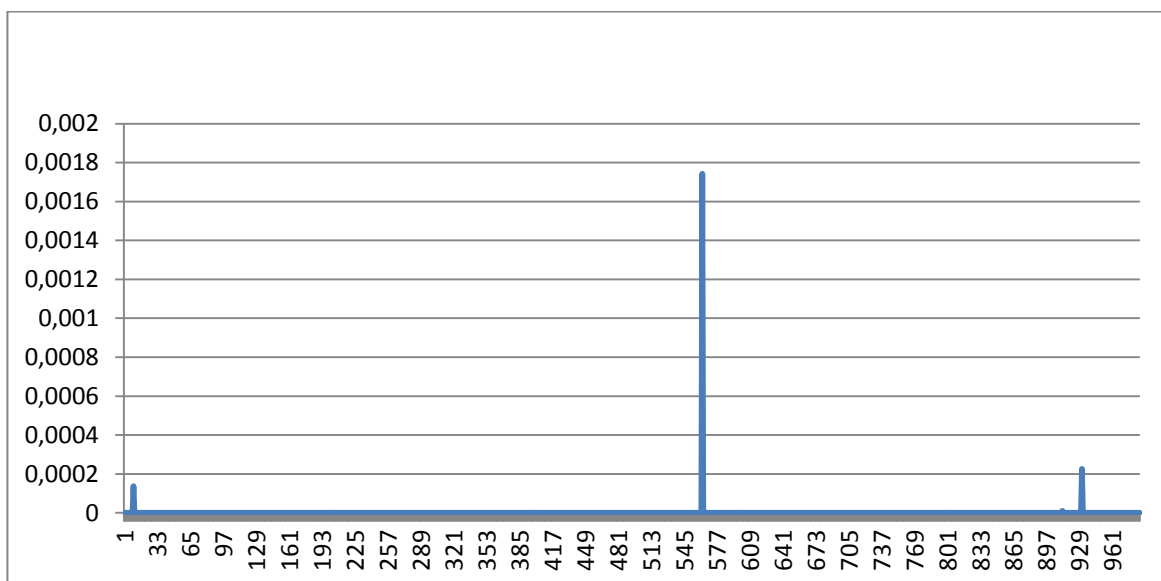


Рисунок 22 – График отклонений показателя корреляции, рассчитанного вручную, и показателя, рассчитанного с помощью разработанного модуля

3.8 Вывод по главе 3

Итогом этой главы является разработанный модуль системы ПО ЭММ. На данный момент для модуля площадного корреляционного анализа реализованы все функции, указанные в требованиях, а также выявлены несколько направлений его оптимизации и развития (указаны в главе 2).

Реализована основная функция — корреляционный анализ площадных данных по заранее заданным парам чисел. Загрузив файл, содержащий таблицу, пользователь может проанализировать его, задав параметры «окна» и шаг, и получить на выходе файл, доступный для анализа с помощью ГИС-пакета.

Впоследствии, планируется разработать несколько более сложных версий модуля согласно рекомендациям заказчика.

4 Анализ экспериментальных данных

Изучение деформационных характеристик земной коры — зон сжатия и растяжения — является важной частью мониторинга процессов геологической среды.

Измерения деформаций на различных временных и пространственных базах позволяют определять движение рудных растворов в геологическом прошлом и предсказывать расположение месторождений в настоящем. Данные о смещениях и деформациях могут быть использованы для уточнения моделей геологической среды, а изучение многолетних характеристик важно для поиска возможных месторождений.

4.1 Теоретическое обоснование

В рамках поставленного эксперимента были изучены данные, полученные на золоторудном участке Североенисейского района в 2014 году. Были поставлены следующие задачи:

- Различить право- и левосторонние сдвиги массивов горных пород,
- Выделить междвиговые области сжатия и растяжения.

Обе задачи взаимосвязаны и могут быть решены на основе анализа гравитационного поля.

В качестве метода решения был выбран гравиструктурный метод, в основе которого лежит более общий метод частотно-азимутального анализа поля, а для изучения гравитационного поля применен традиционный морфоструктурный подход. С позиции этого подхода карта любого геофизического поля является проекцией всех особенностей пространственного распределения того физического свойства горных пород, которым это поле определяется.

Сильная сторона морфоструктурного подхода — его универсальность по отношению к исходным данным, а также большой объем извлекаемой

информации, а слабая – отсутствие строгой аналитической связи источников изучаемого поля с получаемыми результатами и, как следствие, весьма ограниченные возможности постановки и решения прямых геофизических задач [17].

4.2 Результаты эксперимента

В результате эксперимента был проделан площадной корреляционный анализ данных азимутального распределения правых и левых сдвигов деформационной природы на изученной площади, выделенных по удельному электрическому сопротивлению.

Было выявлено, что наиболее крупные месторождения (...) в пределах исследованной площади пространственно тяготеют к областям междвигового растяжения. Причем области растяжения затрагивают, скорее всего, породы нижней части осадочного чехла и кристаллического фундамента, так как для выделения таких областей использованы сдвиги гравитационных аномалий в низкочастотной части спектра.

Кроме этого, в качестве результата следует отметить, что тектонические нарушения, локализованные в непосредственной близости от месторождений, достаточно хорошо коррелируются по простиранию с выделенными областями сжатия/растяжения [18].

4.3 Вывод по главе 4

Итогом данной главы является испытание модуля в рамках реального эксперимента и интерпретация полученных данных для геолого-геофизических целей.

Выявлено, что разработанный модуль корректно справляется со своей задачей и позволяет в полной мере осуществить площадной корреляционный анализ.

ЗАКЛЮЧЕНИЕ

В результате бакалаврской работы выполнены поставленные задачи. Изучена и проанализирована предметная область, построена математическая модель, выявлены и сформулированы требования к разрабатываемому модулю, обоснованы методы и средства разработки, спроектирован, разработан и подготовлен к интеграции с основной системой программный модуль площадного корреляционного анализа геолого-геофизических данных для системы ПО ЭММ.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Разведочная геофизика [Электронный ресурс] // ru.wikipedia.org: Википедия: Свободная энциклопедия. — Режим доступа: http://ru.wikipedia.org/wiki/Разведочная_геофизика
2. Фирчук Ж.В.. Разработка модуля для расчета корреляции между электромагнитными параметрами горных пород геологического разреза и характером рельефа / Ж.В. Фирчук, Д.А. Белоусова // Региональные проблемы дистанционного зондирования Земли: материалы III Междунар. науч. конф. (Красноярск, 13–16 сентября 2016 г.). – 2016. – С. 117–120.
3. Молокова Н. В. Разработка программного модуля для исследования связи свойств геологической среды и рельефа / Н. В. Молокова, Ж. В. Фирчук, Д. А. Белоусова, С. С. Окишев // Молодой ученый. — 2016. — №19. — С. 96-99.
4. Хмелевской, В.К. Геофизические методы исследования земной коры: книга. / В.К. Хмелевской — Международный университет природы, общества и человека "Дубна", Дубна, 1999 г. — 203 с.
5. Комплексирование геофизических методов [Электронный ресурс] // <http://www.mining-enc.ru>: Горная энциклопедия. — Режим доступа : [http://www.mining-enc.ru /k/kompleksirovanie-geofizicheskix-metodov](http://www.mining-enc.ru/k/kompleksirovanie-geofizicheskix-metodov)
6. Бродовой В.В. Комплексирование методов разведочной геофизики. Справочник геофизика: книга. — / В.В. Бродовой, А.А. Никитин, — Недра, Москва, 1981 г. — 397 с.
7. Васнев С.А. Статистика: Учебное пособие — / С.А Васнев — Москва: МГУП, 2001 — 170 с.
8. Корреляционная зависимость [Электронный ресурс] // ru.wikipedia.org: Википедия: Свободная энциклопедия. — Режим доступа: <http://ru.wikipedia.org/wiki/Корреляция>

9. Microsoft Visual Studio [Электронный ресурс] // ru.wikipedia.org: Википедия: Свободная энциклопедия. — Режим доступа: http://ru.wikipedia.org/wiki/Microsoft_Visual_Studio
10. Паттерн MVVM [Электронный ресурс] // professorweb.ru: ProfessorWeb. — Режим доступа: http://professorweb.ru/my/WPF/documents_WPF/level36/36_5.php
11. Белоусова Д.А. Разработка архитектуры программного комплекса для расчета и обработки полевых данных электроразведки / Д.А. Белоусова, С.С. Окишев // Региональные проблемы дистанционного зондирования Земли: материалы III Междунар. науч. конф. (Красноярск, 13–16 сентября 2016 г.) – Красноярск : Сиб. федер. ун-т, 2016. – С. 50-53
12. Иванова И.А. Решение геологических задач с применением программного пакета Surfer: практикум для выполнения учебно-научных работ студентами направления «Прикладная геология» / И.А. Иванова, В.А. Чеканцев. – Томск: Изд-во Томского политехнического университета, 2008. – 92 с
13. CSV [Электронный ресурс] // ru.wikipedia.org: Википедия: Свободная энциклопедия. — Режим доступа: <http://ru.wikipedia.org/wiki/CSV>
14. DAT – Файл данных (Data File) [Электронный ресурс] // fileext.ru: Расширения файлов, типы файлов. — Режим доступа: <http://fileext.ru/dat>
15. Руководство по WPF [Электронный ресурс] // metanit.com: Metanit. Сайт о программировании. — Режим доступа: <https://metanit.com/sharp/wpf/>
16. Трусов П.В. Введение в математическое моделирование: Учебное пособие/ П.В. Трусов. — Москва.: Университетская книга, Логос, 2007. – 440 с.
17. Макеев, С. М. Решение двух задач геодинамики гравиструктурным методом/ С. М. Макеев // Руды и металлы. – 2014. – № 3. – С. 63–74
18. Макеев, С. М. Пространственная корреляция гравитационного поля и золоторудных объектов Енисейского кряжа/ С. М. Макеев // Руды и металлы. – 2014. – № 3. – С. 17–26